1

DECISION MAKING IN THE FACE OF UNCERTAINTY—I (Uncertain Outcome)

Richard Bellman

P-568

21 September 1954

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

Summary: ~~In this paper we consider~~ a number of simple multi-stage decision processes where the intuitive concept of maximizing expected gain over expected cost is valid are considered.

DECISION MAKING IN THE FACE OF UNCERTAINTY—I
(Uncertain Outcome)

Richard Bellman

§1. Introduction

In logistics there is a large class of situations in which we are to use a given resource repeatedly until it is exhausted. If at each stage we have several different ways of utilizing this resource, it is a problem of importance to determine the procedure which maximizes the over-all value of the resource. In a previous paper [1] we have given some applications of the theory of dynamic programming to problems of this general class.

In this paper we wish to consider some representative problems where the outcome of any stage is not completely determined. We shall show that in some cases optimal policies of quite simple and intuitive nature exist.

In subsequent papers we shall discuss some cases of uncertain purpose, which is to say, processes in which we do not completely know the form of the "payoff" or criterion function.

§2. Discussion

A problem of recurrent type is the following: We have a certain quantity of resource which we can use in various fashions. Any particular application yields a certain distribution of returns balanced against a certain distribution of costs. How do we proceed to use the resource until depleted so as to maximize the average return?

A naive approach to the problem runs as follows: As a result of any particular decision we obtain a certain expected gain and suffer a certain expected cost. A reasonable procedure for maximizing over-all gain would then seem to be one which yields the maximum return per unit cost, i.e., one which maximizes the ratio

$$R = \frac{\text{immediate expected gain}}{\text{immediate expected cost}} \tag{2.1}$$

This prescription for an optimal policy has many desirable features. It is simple, intuitive, requires only an estimation of average values rather than detailed knowledge of the distribution of events—and is occasionally correct.

We shall discuss below a number of situations in which it is approximately correct.

The purpose of the simple models discussed below is to bolster our intuition, which is, after all, only a combination of the results of theory and practice. As far as the multi-stage processes of realistic type are concerned, existent theory is meager. Consequently, it is important to build up a backlog of as many mathematical models as possible with known solutions in order to obtain clues to the realistic problems which usually defy precise analysis.

§3. A Multi-stage Allocation Problem of Deterministic Type

Let us begin with a simpler process which is deterministic. We have an initial resource x which may be utilized in a number of ways. If y is a parameter specifying a particular use, we let $R(x,y)$ be the immediate return, and $D(x,y)$ the cost in resources. How do we proceed to utilize this resource so as to maximize the total return?

We set (see [1], [2], [3])

$$f(x) = \text{total return from an initial resource } x, \text{ using an optimal allocation policy} \tag{3.1}$$

Then, as discussed in [1], [2], this function satisfies the functional equation

$$f(x) = \operatorname*{Max}_{y} \left[R(x,y) + f(x - D(x,y)) \right] \tag{3.2}$$

Let us now make the fundamental assumption that $D(x,y)$ is small compared to x for any y, i.e. $0 < D(x,y) << x$. Proceeding formally, we may write

$$f(x) = \operatorname*{Max}_{y} \left[R(x,y) + f(x) - D(x,y)f'(x) + \cdots \right] \tag{3.3}$$

which yields the approximate equation

$$0 = \operatorname*{Max}_{y} \left[R(x,y) - D(x,y)f'(x) \right] \tag{3.4}$$

This means that for one y, say $\bar{y}$, we have

$$0 = R(x,\bar{y}) - D(x,\bar{y})f'(x) \tag{3.5}$$

and for all others we have

$$0 \geq R(x,y) - D(x,y)f'(x) \tag{3.6}$$

Consequently, within the error contained in using (3.4) in place of (3.3), we have

$$f'(x) = \operatorname*{Max}_{y} \frac{R(x,y)}{D(x,y)} \tag{3.7}$$

which is equivalent to the statement that at each stage we use our resource in accord with the prescription of (1.1).

§4. Multi-stage Allocation—Stochastic

Let us now consider the process above where we have a distribution of returns and allocations. Any particular utilization yields a set of returns, z, characterized by a distribution function dR(y,z,s), and a set of costs, w, characterized by a distribution function dD(y,w,x). We now wish to maximize the expected total return from an initial resource x. Denote this quantity by f(x). Then, as above, f(x) satisfies the equation

$$f(x) = \operatorname*{Max}_{y} \left[\int_0^\infty z\,dR(y,z,x) + f\left(x - \int_0^\infty w\,dD(y,w,x)\right)\right] \tag{4.1}$$

Assuming, as before, that

$$\text{Expected Cost} = \int_0^\infty w dD(y,w,x) \ll x \tag{4.2}$$

we obtain as an approximation to (4.1), the equation

$$0 = \underset{y}{\text{Max}} \left[\int_0^\infty z d(R,y,z,x) - f'(x) \int_0^\infty w dD(y,w,x) \right] \tag{4.3}$$

or

$$\begin{aligned} f'(x) &= \underset{y}{\text{Max}} \int_0^\infty z dR \Big/ \int_0^\infty w dD \\ &= \underset{y}{\text{Max}} \frac{\text{Immediate Expected Return}}{\text{Immediate Expected Cost}} \end{aligned} \tag{4.4}$$

which is precisely the prescription of (1.1).

§5. Conclusion

We have discussed above two representative examples of multi-stage allocation processes, in each of which the prescription of (1.1) furnished an approximation to the optimal policy.

In those situations where the change in resources may be large compared to initial resources, the above analysis is not as useful. Nonetheless, even in these cases, this prescription furnishes a useful first approximation which may be improved by successive approximations based upon the functional equations (see [1], [2], [3]).

BIBLIOGRAPHY

1. Bellman, R., "On Some Applications of the Theory of Dynamic Programming to Logistics," Naval Research Logistics Quarterly (to appear).

2. ————, "Some Problems in the Theory of Dynamic Programming," Econometrica, Vol. 22, No. 1 (January 1954), pp. 37–48.

3. ————, "Some Applications of the Theory of Dynamic Programming—A Review," Journal of the Operations Research Society of America, Vol. 2, No. 3 (August 1954), pp. 275–288.

bj